Journal of Marine Science and Technology
https://doi.org/10.1007/s00773-024-01043-x

ORIGINAL ARTICLE

Improving decision transparency in autonomous maritime collision avoidance

Andreas Nygard Madsen[1] · Andreas Brandsæter[2,3] · Koen van de Merwe[4] · Jooyoung Park[5]

Received: 30 April 2024 / Accepted: 8 December 2024

Abstract

Recent advances in artificial intelligence (AI) have laid the foundation for developing a sophisticated collision avoidance system for use in maritime autonomous surface ships, potentially enhancing maritime safety and decreasing the navigator's workload. Understanding the reasoning behind an AI system is inherently difficult. To help the human operator understand what the AI system is doing and its reasoning, we employed a human-centered design approach to develop transparency layers that visualize different aspects of an operation by displaying labels, diagrams, and simulations intended to improve the user's situation awareness (SA). The effectiveness and usability of the different layers were investigated through simulator-based experiments involving nautical students and licensed navigators. The SA global assessment technique was utilized to measure navigators' SA. User satisfaction was also measured, and effective layers were identified. The results indicate that the transparency layers that enhance SA Level 3 are preferred by participants, suggesting a potential for improving human–AI compatibility. However, the introduction of transparency layers does not uniformly enhance SA across all levels, and a tendency toward passive decision-making was observed. The findings highlight the importance of balancing information presentation with the user's cognitive capabilities and suggest that further research is needed to refine transparency layers for optimized human–AI compatibility in maritime navigation.

Keywords Autonomous Ships · MASS · Human machine collaboration · Decision transparency · Decision support · Simulator testing · HMI · Humans in the loop · Shore control center

1 Introduction

Early on the morning of November 8, 2018, the frigate Helge Ingstad collided with the tanker Sola TS in the Norwegian Hjeltefjord. The crew on board the Helge Ingstad was led to believe that the approaching tanker was an oil terminal due to the deck lights used for tugboat operation. Even though the automatic identification system (AIS) and the automatic radar plotting aid identified the tanker and the Sola TS contacted the Helge Ingstad through VHF radio, the crew was not able to determine the risk of collision and take necessary action before it was too late. According to Part 1 of the investigation report (AIBN & DAIBN, 2019), inadequate situation awareness (SA) was the primary factor in the accident. Regarding the Helge Ingstad, the report refers to poor individual SA and states that systemic factors, such as procedures and design, were not optimized to ensure good SA. Combined with the officer on watch's certainty in his own SA (AIBN & DAIBN, 2019, p. 143), this led to an inability to rectify erroneous SA [1].

With recent advances in artificial intelligence (AI) and autonomous ships, the future might bring forth ships that are equipped with systems, such as an autonomous collision avoidance system (CAS), capable of both decision support and decision-making. In the Helge Ingstad case, an autonomous CAS could have performed a collision avoidance maneuver, for example, reducing the frigate's speed

✉ Andreas Nygard Madsen
andreas.madsen@ntnu.no

1 Department of Ocean Operations and Civil Engineering, Norwegian University of Science and Technology, Ålesund, Norway

2 Department of ICT and Natural Sciences, Norwegian University of Science and Technology, Ålesund, Norway

3 Department of Science and Mathematics, Volda University College, Volda, Norway

4 Group Research and Development, DNV, Høvik, Norway

5 Department of Design, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Published online: 06 January 2025

or making it come to a full stop. However, how introducing such a system will affect the crew's SA is an open question.

The International Maritime Organization [2] has defined four degrees of autonomy in its regulatory scoping exercise for autonomous ships. These levels extend from manned ships equipped with automation for certain processes and decision support to completely autonomous (unmanned) ships that can make decisions and act independently through their operating systems. Having an autonomous CAS approved for operation with no human supervision brings with it major regulatory challenges. Additionally, such a system could perhaps be at its best when collaborating with a human. Therefore, the focus of this paper is on the type of tasks the human navigator shall embark on when operating or supervising a ship, without emphasizing the physical location of the human operator.

To achieve successful human–AI cooperation, the human navigator would need to understand the system's decisions and be capable of intervening when (and only when) needed. Previous research has demonstrated that this can be quite challenging when an autonomous CAS does not explain its decisions [3, 4]. This is also in line with research in other domains, such as aviation [5].

To address this gap, the present study introduces a layered approach to transparency that explicitly visualizes the CAS's decision-making process. This approach provides navigators with tiered insights into the why behind system actions, facilitating a deeper understanding and, consequently, more informed and timely intervention when needed. By enhancing individual SA and making the CAS's decisions more interpretable, the layered transparency approach is intended to increase human–AI compatibility, ultimately promoting smoother, more effective cooperation between human navigators and autonomous systems in maritime navigation. Transparency has been defined as enhancing the understandability and predictability of systems [6], and recent research has found positive effects on both human SA and performance when the transparency of systems is increased (van [7]). This is supported by van de Merwe et al. [8], who found that increased transparency led to increased SA and performance during collision avoidance tasks in an experimental trial in which the participants interpreted still images. Their results also indicate that users prefer higher levels of transparency than lower levels.

Here, the transparency layers were developed through a human-centered design (HCD) process with four phases [9], and they were tested with navigators in a full-mission simulator. The HCD approach prioritizes understanding the needs, behaviors, and requirements of the people who will be using the system being designed. It focuses on the end users throughout the design process, with the aim of creating solutions that are intuitive, user-friendly, and tailored to meet the specific needs of those users [10].

Based on Endsley's [11] three-level model for SA, treating the AI's decision as a product of these three levels can be meaningful. By allowing the navigator to tap into the system's SA, represented by a transparency layer for each SA level, the AI's decisions will have increased transparency. The first phase of the HCD process was to understand and specify the context of use, which included a literature review of how decision transparency for collision avoidance has been treated in the existing literature [12] and an investigation into deviations from the International Regulations for Preventing Collisions at Sea (COLREG) and ship interactions [13]. The second phase, specifying the user requirements, was to develop a simulator-based approach to test and assess human performance when tasked with supervising an autonomous CAS [3] and to discover what kinds of transparency are called for by the users [4]. The third phase was to develop a design solution for visualizing the transparency layer (Sect. 3.2), while the fourth phase was to evaluate the design presented in this paper. When evaluating a solution, the focus is on systematically assessing its usability. According to the ISO standard [9], this includes effectiveness, effectivity, and user satisfaction.

The challenge lies in ensuring human operators' trust and comprehension of AI-driven CAS, where opacity in decision-making can impede effective human–AI collaboration. This research investigates transparency layers as a solution, examining how structured information presentation affects navigators' SA and decision-making.

The main contribution of this work is therefore that it investigates how the introduction of transparency layers for an AI's decisions influences a navigator's individual SA. This was answered through the following research questions (RQs):

- RQ1: How do the different layers contribute to the navigator's acquisition of individual SA?
- RQ2: Which layer is the most effective?
- RQ3: Which layer provides the best user satisfaction?

This paper proceeds by first detailing the experimental setup and methodology, followed by an analysis of results across SA, decision effectiveness, and user satisfaction. The discussion section addresses the implications of these findings, concluding with directions for future research.

2 Background

Technological advancements make autonomous ships, which are navigated primarily by AI, a possibility in the future. This represents a leap forward in maritime navigation, but it also introduces complex challenges, particularly for navigators tasked with supervising these vessels. Understanding the implications of supervising an autonomous ship and the difficulties in maintaining SA is important for ensuring safe and efficient maritime operations [13–15]. The navigator's tasks have been changed multiple times before,for example, both radar and the electronic chart display and information system (ECDIS) were disruptive innovations, changing the navigation practice and reducing the navigator's workload. However, along with the introduction of these systems, new errors occurred, such as ECIDS-assisted accidents [16] and radar-assisted accidents [17]. The lack of a human-centered approach when developing the ECDIS has been identified as a root cause of many incidents [18–20], and learning from history, a human-centered approach toward autonomous ships may hold the key to avoiding or at least reducing, autonomy-assisted accidents in maritime collisions.

When engaged in supervisory tasks, where the system has high reliability, humans tend to not supervise properly, which is known as automation complacency. Another possible consequence of automation complacency is what is known as an out-of-the-loop syndrome [21], which is closely linked to SA [22]. Out-of-the-loop syndrome describes a situation in which operators become disengaged or lose SA due to excessive reliance on automation [23]. As automation takes over more tasks, operators may become less involved in the actual control or monitoring of the system, which leads to a decreased ability to intervene effectively in the event of automation failure or unexpected situations [22].

When supervising an autonomous CAS, we anticipate that the navigator will frequently be out of the loop, and if the AI-based system calls for the navigator's attention, they will not necessarily have adequate SA and might struggle to acquire and maintain adequate SA in a timely manner.

2.1 Situation awareness (SA)

One of the most used models to explain SA is Endsley's three-level model [11]. According to Endsley and Jones [24], the construct of SA can be summarized as "being aware of what is happening around you (level 1), and the understanding of what that information means to you now (level 2) and in the future (level 3). (p. 13)" This cognitive theory uses an information-processing approach, and the three levels are displayed in Fig. 1.

Wickens [25] argues that SA has a functional value and that the focus should not be on proving or disproving a particular model for SA: "Allowing a certain fuzziness enables concentration to be redirected away from proving right or wrong toward the utility of the concept in applications" (p. 90). With regard to a navigator operating or supervising an autonomous ship, measuring the navigator's SA has a functional value, and we will allow here a certain fuzziness.

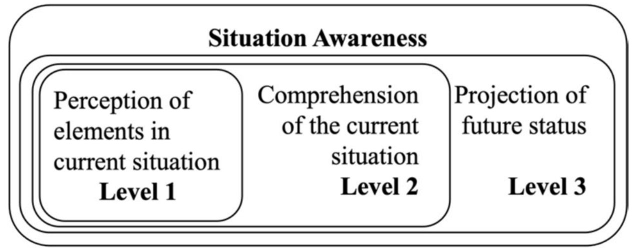


Fig. 1 Three-level model for SA [11]

Endsley and Jones [26] argue for shared SA in teams, in which each member of a team holds a unique individual SA and, through design, one can make sure that parts of individual SA overlap with those of the other members of the team so that all team members have a shared SA regarding the operations. Stanton et al. [27] argue for the concept of compatible SA, in which no individuals working within a collaborative system hold the same SA, nor is it a goal to achieve this. The main objective is to ensure that the SA of each human (agent) within the system is different in content but still compatible since a collective awareness is required for the system to successfully collaborate in solving tasks. The compatible SA model is illustrated as a puzzle in which all team members are compatible with each other.

The need for explanations varies among stakeholders [28], and the navigator does not need to know everything about the AI's algorithmic calculations or everything it has detected or computed. Humans must construct their own understanding of the AI [5], and explanations need to be tailored to the situational circumstances [29]. By the same token, the AI for collision avoidance does not need to know the human's complete mental model, but it will need to understand certain aspects of human behavior, such as deviations from collision regulations [13], to avoid situations that may reduce safety and efficiency of operations. When considering human-–AI interaction, in which both the human and the AI are tasked with decision-making, it is reasonable to treat the design process as a need for compatible SA rather than shared SA.

In this article, we focus on the compatibility between an AI that performs autonomous navigation and the navigator who supervises the navigation. Here, we propose that the AI (artifact) can "explain" its decisions through transparency layers based on Endsley's three-level model

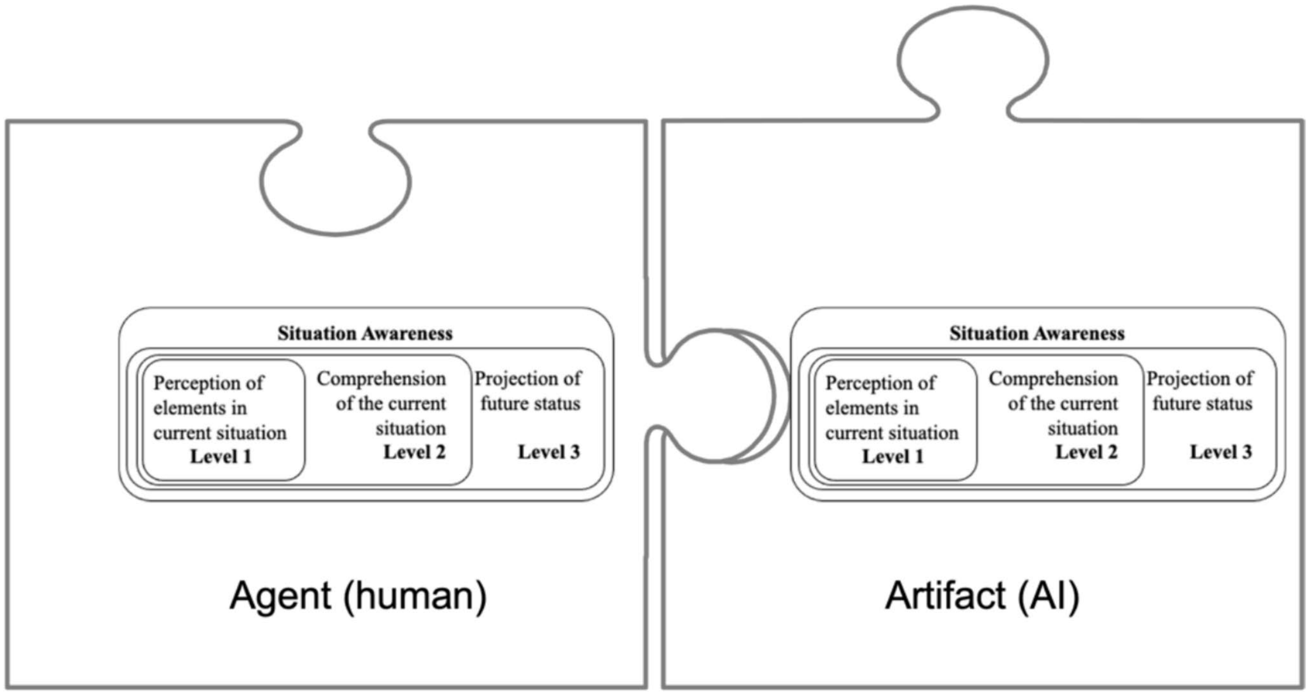


Fig. 2 Agent–Artifact combability with Endsley's three-level model [11]

[11] to increase compatibility with the navigator (agent): (1) what the AI has perceived, (2) how it comprehends the perceived information, and (3) what it believes will happen in the near future (Fig. 2).

By utilizing the SA global assessment tool (SAGAT) to measure how the introduction of the transparency layers affects the navigator's SA, this study can provide a simpler and more intuitive description of SA.

3 Methods

This section describes how the study was designed and the development of the transparency layers, and it explains the methods utilized for each RQ, as summarized in Table 1.

Based on the RQs, null hypotheses were formulated to test the results.

The dependent variables in Table 1 were selected based on the ISO standard 9241–210:2010 for HCD, which describes the evaluations of designs. This stage focuses on assessing the effectiveness, usability, and satisfaction of the design solutions developed to ensure that they meet users' needs and achieve the desired outcomes [9].

3.1 Simulator experiment

The research was conducted using K-Sim full-mission bridge simulators (Kongsberg Digital, Norway), normally used for deck officer training. The setup included two identical simulation bridges, each equipped with three visual displays for the forward view and one for the aft view. Additionally,

Table 1 Research questions, variables, method, and hypotheses

Research question	Variables		Method	Hypothesis
	Independent	Dependent		
RQ1	Transparency layers	Situation awareness	SAGAT	H_{01} –The three transparency layers developed in this research do not contribute to navigators' SA
RQ2	Transparency layers	Effectiveness (decision)	Confusion matrix	H_{02}- The three transparency layers developed in this research do not affect how effective the navigator is
RQ3	Transparency layers	Satisfaction	Card-rating technique	H_{03} – There is no difference in preference between transparency layer 3 and other levels of transparency

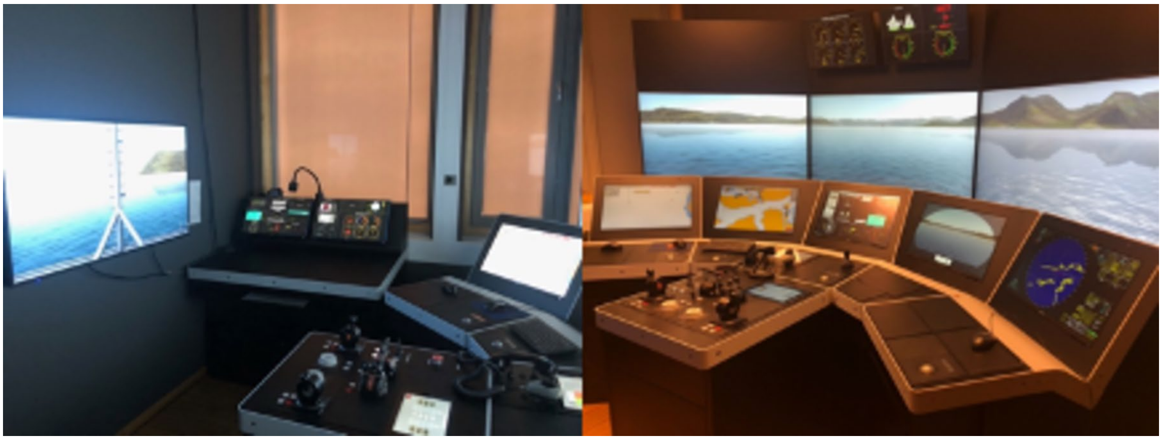


Fig. 3 K-Sim full-mission bridge simulator, aft and forward

each simulator featured a binocular screen with a joystick for panoramic observation, dual ECDISs, a conning display, radar, and various control panels. The hardware setup also encompassed devices for steering, speed control, and radio communication (Fig. 3).

The participants were briefed on their involvement in the research project. They were informed that the simulator would feature an autonomous CAS executing evasive maneuvers that would sometimes call for human intervention. The students were provided with detailed information about the ship and the geographic context of each scenario.

Participants were directed to enter the bridge simulator when a monitored ship requested human supervision. The task of the participant was to supervise the system's decision. After carefully assessing the scenario, the participants decided whether it was necessary to intervene and take manual control of the system, aborting the maneuver proposed by the system. When they made their decision to intervene or not, they exited the bridge, and the simulation stopped.

3.2 Independent variable: transparency layers

To test how transparency layers may affect navigation, we created a mock-up AI-based autonomous CAS. Before the experiment, all participants took part in a familiarization exercise in which they tested all four transparency layers and were informed that they would randomly have only one of the layers available in each scenario.

The mock-up was created by making screen recordings from an ECDIS. The CAS's decision was added to the ECDIS as a sailing route, and each of the four scenarios was played through with screen recording two times—one in which the ship followed the AI's decision, and one in which the ship maintained the course and speed. Using Adobe After Effects, each video was edited to add specific visual layers for transparency. As the SA level advanced, additional layers were accumulated. This made it possible to run the mock-up as a Wizard of Oz solution, without making the participants aware that it was a mock-up.

The transparency layer was created to ensure compatibility between the navigator and the CAS. Based on the findings from the literature review [12] and the user feedback [4], visualization of the transparency layers was developed in collaboration between the authors. The idea was to utilize Endsley's three-level model of SA to make the AI's SA transparent. The CAS would then provide information about its perception (Layer 1), comprehension

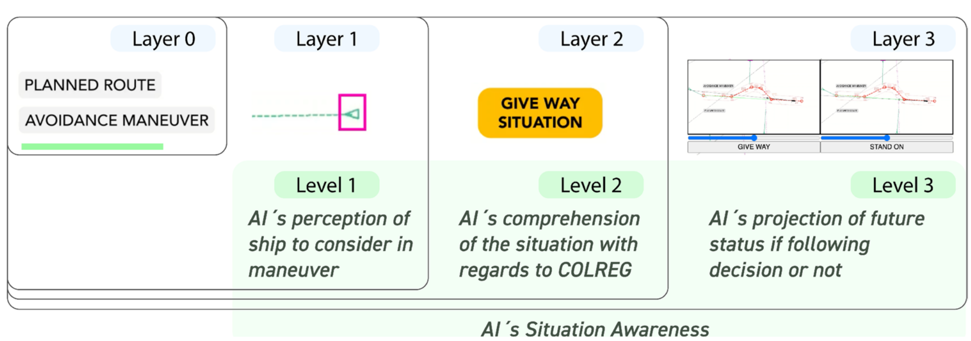


Fig. 4 Visual features for transparency layers were added to the video according to SA levels

(Layer 2), and projection of the future state (Layer 3) (Fig. 4).

Transparency layer 0 displays no information from the CAS's SA, only the decision. Labels for "planned route" and "avoidance maneuver" are added to indicate the trajectories where the ship maintains the original route or follows the CAS's decision. The original route is marked with a green line for better visibility among the other lines on the ECDIS. In some situations, this might be sufficient for the navigator to understand why the CAS made a certain decision, since the decision itself can be self-explanatory to a trained navigator, at least in situations with low complexity.

Transparency layer 1 represents the system's perception. The ship(s) that has been considered in a COLREG maneuver is highlighted by a pink rectangle (bounding box). This layer explains which ship(s) has been considered in an assessment; hence, vessels without a bounding box have not been considered in the assessment.

Transparency layer 2 represents how the CAS comprehends the situation, and signage for the "give-way situation" is added to display the AI's decision. The opacity of the text fluctuates between 100 and 0, creating a blinking effect to differentiate it from the labels added at Level 0.

Transparency layer 3 illustrates the system's projection of the future status. Here, the navigator is allowed to use a trial function, where they can fast-forward how the AI believes the situation will develop when performing a give-way or stand-on maneuver. Two small video screens are added to project the expected near future in fast-forward. By clicking the buttons underneath, the navigator can play and watch full videos of both give-way and stand-on maneuvers. The videos were edited to play 5 times faster, and the average play time of each small video is thus around 5 s (Fig. 5).

Each of the main videos (Scenarios 1–4, Levels 0–3) has a running time of 2 min and showcases the movement of the ownship from the starting point of the trajectory to the point where the avoidance maneuver commences.

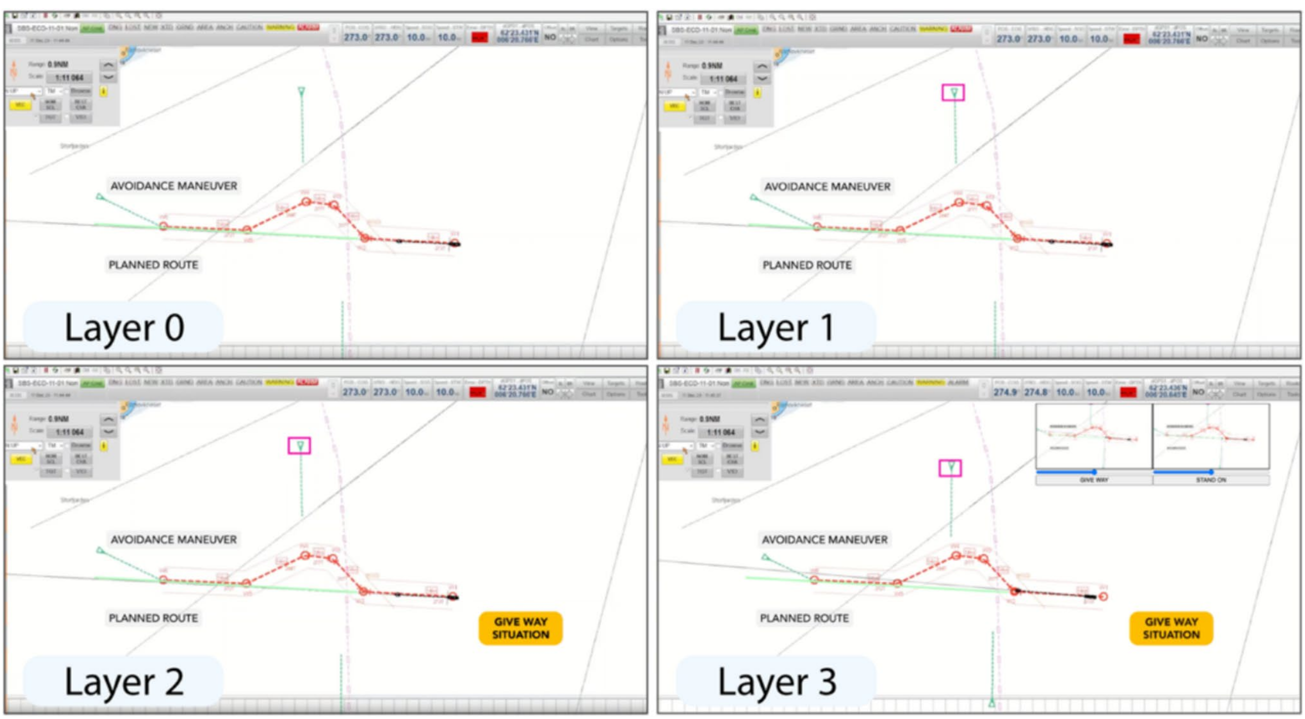


Fig. 5 Screenshot of the familiarization scenarios with each transparency layer

Table 2 Participants' age, experience at sea, education, and type of certificate

Means ± standard deviations							
Age		Sea time		Education		Certificate	
Students	27.8 ± 7 yrs	Students	8.3 ± 7 yrs	Vocational:	$n=9$	D1:	$n=11$
Certified	42.7 ± 13 yrs	Certified	11.4 ± 9 yrs	University College:	$n=4$	D2:	$n=4$
Total	39.7 ± 13 yrs	Total	10.8 ± 9 yrs	University:	$n=7$	D3:	$n=1$

3.3 Participants

The participants were four nautical students enrolled in vocational education and 16 licensed navigators. The participants were selected by availability and can be viewed in Table 2.

The highest deck officer certificate held by the participants in this study was a master mariner certificate (D1), which allows the person to be a master onboard any kind of ship. Next is D2, followed by D3. The four students did not yet have certificates.

3.4 Randomization

Each participant contributed to exactly four experiments. In each experiment, the participants were randomly assigned a scenario and a transparency layer (random draw, without replacement). The total number of participants per combination of the transparency layer and scenario is shown in Table 3.

Since the scenarios and layers were randomly drawn, the number of experiments in each combination varied. Therefore, when comparing the results, we limited our analysis to summary statistics for each combination, such as the mean, median, standard deviation, and so on. The varying number of experiments in each combination of transparency layer and scenario may have influenced the results.

3.5 Scenarios

Based on real situations identified in a set of radar and AIS recordings at a ferry crossing in Storfjorden, Norway [13, 30], four scenarios were created by the main author, who is a licensed mariner. The simulator scenarios are four of the eight scenarios presented in Brandsæter & Madsen [3].

For each scenario, a collision avoidance maneuver for the CAS was created by the main author. Two of the scenarios include situations in which the CAS intentionally proposes suboptimal maneuvers. Here, the correct action for the human supervisors is to intervene. In the other two scenarios, the proposed maneuver is considered optimal, and the correct intervention is to not intervene. All scenarios were approved by three independent subject matter experts (SMEs), all of whom agreed on the preferred action.

All scenarios are described in detail in Appendix I. Here, we use Scenario 1 as an example and provide an explanation of the scenario's complexity and the experts' opinions of the solution provided by the mock-up CAS. In Fig. 6, the ownship is the vessel marked with a black double circle or ship outline, and the target vessels are the green triangles (AIS symbols).

Table 3 Participants in each scenario per layer

Scenario 1		Scenario 2		Scenario 3		Scenario 4	
Layer 0	8	Layer 0	5	Layer 0	5	Layer 0	2
Layer 1	5	Layer 1	6	Layer 1	6	Layer 1	3
Layer 2	4	Layer 2	4	Layer 2	5	Layer 2	7
Layer 3	3	Layer 3	5	Layer 3	4	Layer 3	8
Sum	20	Sum	20	Sum	20	Sum	20

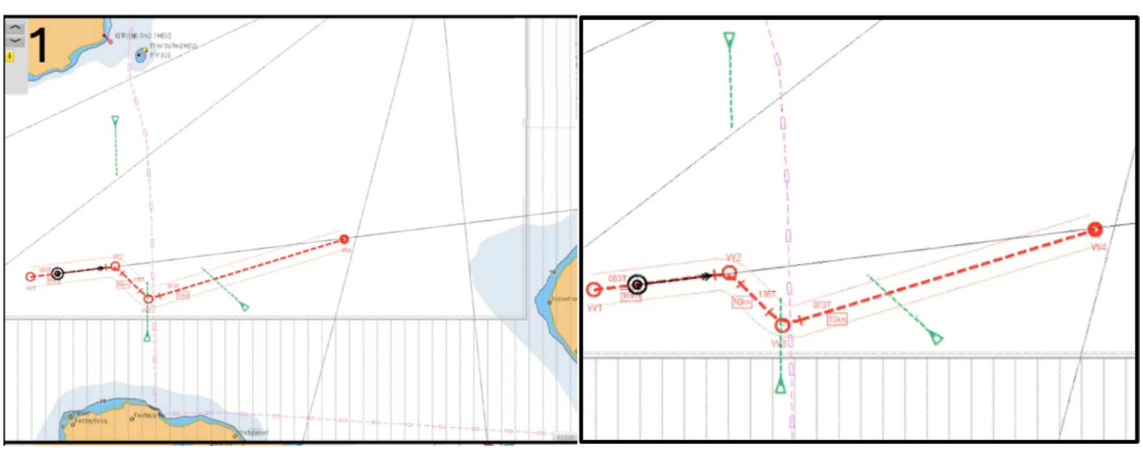


Fig. 6 Screenshot from the starting point in Scenario 1 used in this study. On the right, the scenario is scaled up

In this scenario, the ship heading west-northwest does not appear to be a collision risk at this stage. Either way, the ownship should give way to this ship since this is a crossing situation according to the COLREG. The system's decision appears to consider this and gives way to both ships approaching from the starboard side. However, as the scenario evolves, the ship heading west-northwest changes its course toward the west, creating a head-on situation. Thus, the maneuver decision by the system is not sufficient, and the turn toward starboard must start earlier, meaning that the correct action for the navigator would be to intervene.

3.6 Dependent variables

3.6.1 Situation awareness (SA)

The SAGAT [31] stands as one of the most extensively utilized methods for evaluating SA. It offers an impartial measure of SA and was formulated to measure SA objectively, covering all its facets based on an assessment of the operator's SA needs [32]. When employing the SAGAT, simulations of typical tasks or scenarios are halted at random intervals, and system displays are obscured while participants respond to queries regarding their current understanding of the situation. The participants' responses are then compared with the actual situation. This evaluation serves as a foundation for an impartial assessment of SA. The SAGAT involves queries across the three SA levels. By probing across the entirety of an individual's SA, the SAGAT mitigates potential biases in attention, as candidates cannot anticipate the questions in advance [32].

Due to the short length of the scenarios in this study, there was no "stop and probe" during the simulations, but the participants were queried immediately after each scenario. The queries are described in Fig. 7, and the participants answered the queries immediately after each of the four scenarios.

After completing the SAGAT form, the participants were asked to rate how certain they were of their understanding of the situation on a visual analog linear scale ranging from uncertain to certain, resulting in the variable *certitude* ranging between 0 = uncertain and 100 = certain.

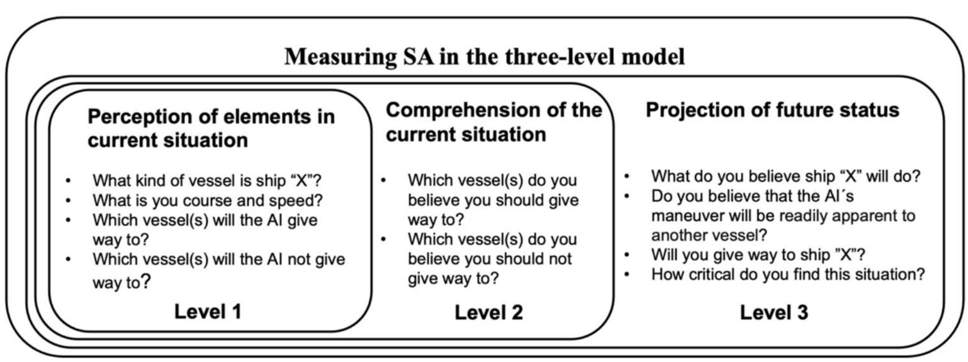
3.6.2 Effectiveness (confusion matrix)

The task measured was the navigator's performance and capability to intervene with an AI system when (and only when) appropriate. By defining the correct decision for each scenario, it was possible to determine whether the participant's decision was right or wrong. The results from the experiments in the simulator-based approach can be presented in a confusion matrix, making it easier to see when a model is confusing one class for another [33]. Although traditionally used for assessing algorithms in supervised learning, the principles of a confusion matrix can be applied to measure human decision-making performance in situations that require active or passive decisions, assuming we know the correct action.

When evaluating the navigators' performance, the decisions were classified as either correct ("true") or incorrect ("false") when compared to the predefined scenarios. Additionally, we categorized decisions based on the nature of the action taken: "positive" for active decisions, such as choosing to intervene with the autonomous CAS's decision, and "negative" for passive decisions, such as choosing not to intervene with the AI's judgment.

- True positives: Instances in which a person correctly decided to take action. (The navigator did intervene, and this was the preferred action.)
- True negatives: Instances in which a person correctly decided against taking action. (The navigator did not intervene, and this was the preferred action.)
- False positives: Instances in which a person incorrectly decided to take action. (The navigator did intervene, but this was not necessary or even dangerous.)

Fig. 7 Queries for each SA level



- False negatives: Instances in which a person incorrectly decided against taking action. (The navigator did not intervene but should have.)

This framework allowed us to assess human decision-making in contexts in which their actions directly influenced outcomes, providing insights into their judgment and intervention effectiveness.

The ideal behavior of the navigator is to intervene when needed and not intervene when not needed. If doing so, only true positive and true negative actions were recorded, leading to a sensitivity and a specificity of 1.

Furthermore, participants who decided to intervene were asked to justify why they intervened with the system.

3.6.3 User satisfaction

Various techniques involving cards have been applied in the design of computer systems, and the advantages of such techniques are well-reported [34]. In this paper, we utilized a card-ranking technique, in which each transparency layer was illustrated on different cards. After completing the four scenarios, the participants were asked to put the cards in order, from the one they preferred the most to the one they preferred the least.

4 Results

The following section provides the results for each RQ.

4.1 RQ1: How do the different layers contribute to the navigator's acquisition of individual SA?

Although the transparency layers are meant to coincide with each respective SA level, viewing the effect of all transparency layers on each SA level can be useful. If a participant has access to a specific transparency layer (x), then the participant also has access to the preceding transparency layers ($x-1, x-2, \dots, 0$). Thus, with access to Layer 3, access to Layers 0, 1, and 2 is also given.

4.1.1 Level 1 SA perception

For SA Level 1, there was no indication of a positive effect of transparency layers. On the contrary, SA Level 1 seemed to decrease when transparency layers were added, especially in Scenarios 3 and 4. A large variation can be observed between the navigators' SA scores in each scenario and within each transparency layer, as illustrated in the box plot. (Fig. 8)

4.1.2 Level 2 SA comprehension

Transparency Layer 2 provided the lowest SA Level 2 score across all scenarios, except Scenario 2, which provided the second lowest score. Layer 2 did not seem to provide a satisfactory Level 2 SA (only 72.2). Transparency Layer 3 was the layer that provided the overall highest Level 2 SA score (Fig. 9).

4.1.3 SA Level 3 projection

For SA Level 3, transparency layer 3 was the layer that provided the highest SA score, as illustrated in Fig. 10. Layer 3

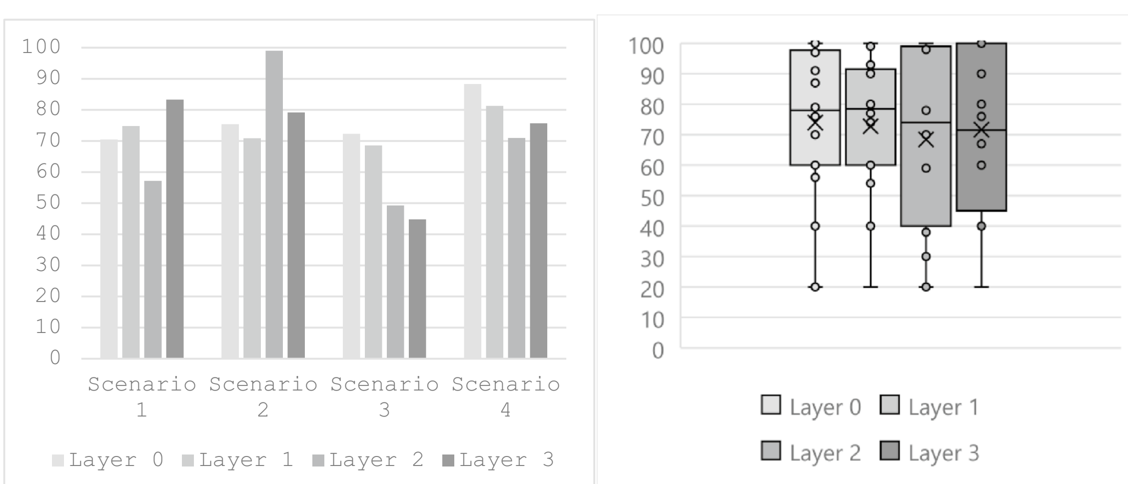


Fig. 8 The bar chart illustrates the mean Level 1 SA score in each scenario for each transparency layer. The box plot shows the overall mean score (X) for each transparency layer and the distribution of the data. Error bars represent the 95% confidence interval

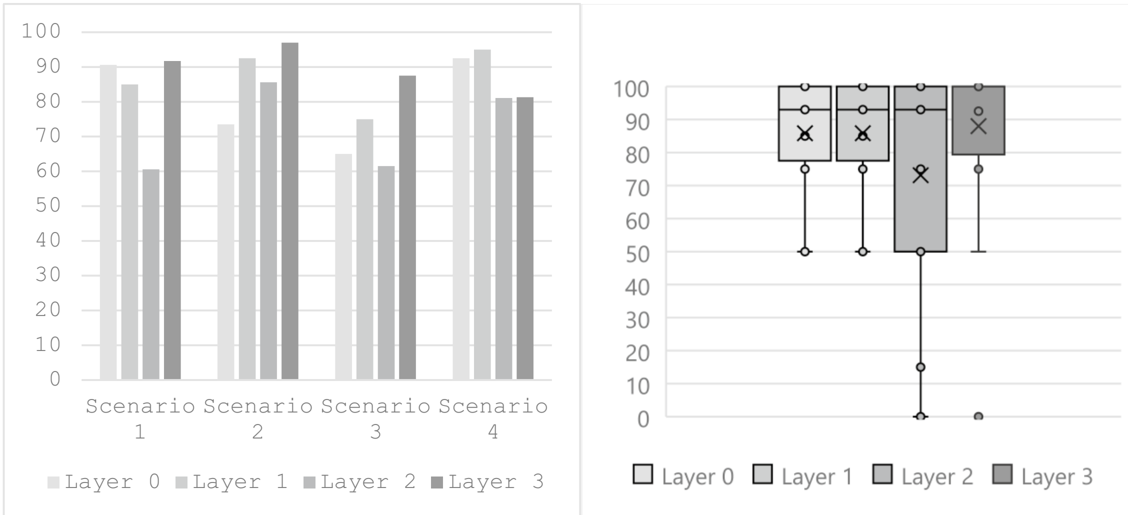


Fig. 9 The bar chart illustrates the mean SA Level 2 score in each scenario for each transparency layer. The box plot shows the overall mean score (X) for each transparency layer and the distribution of the data. Error bars represent the 95% confidence interval

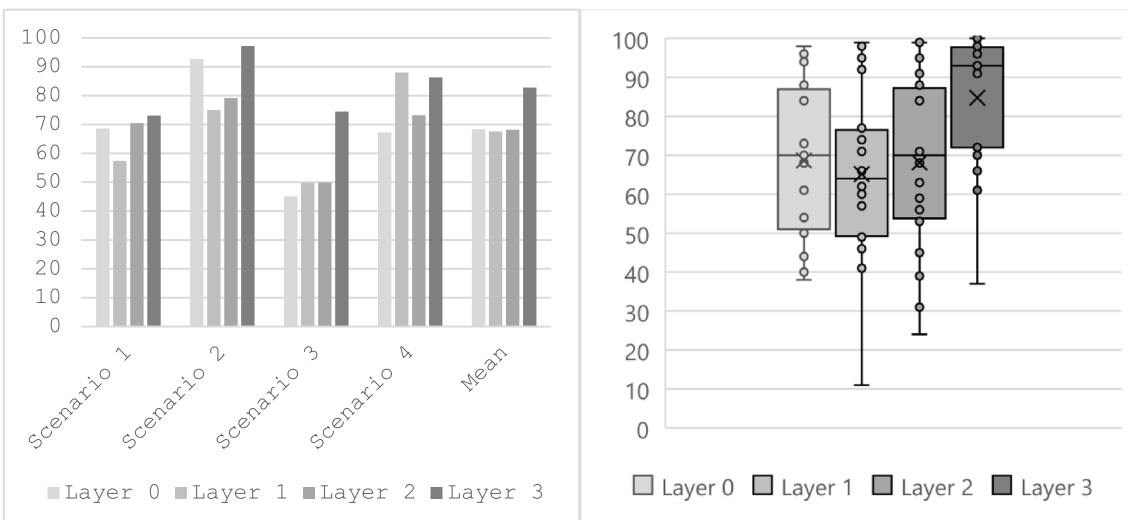


Fig. 10 The bar chart illustrates the mean SA Level 3 score in each scenario for each transparency layer. The box plot shows the overall mean score (X) for each transparency layer and the distribution of the data. Error bars represent the 95% confidence interval

consistently provided high scores across all scenarios. The box plot seems to indicate that there was less variance in SA scores between the navigators who utilized this layer and those who utilized other layers. A closer examination of variance within each scenario is provided in Fig. 11.

Each scenario had a large spread in SA scores. Especially for Scenario 3, much higher SA scores were achieved for the navigators who utilized Layer 3; however, some still had low scores.

4.1.4 Three Levels of SA

For an overview, the mean score for all three levels of SA within each transparency layer is compared in Fig. 12.

Here, transparency layer 3 appears to be equally good, or perhaps slightly worse, at supporting SA Level 1. For SA Level 2, a generally higher score can be seen than for SA Levels 1 and 3, and transparency layer 3 provided the best score. Transparency layer 3 was also better than the others at supporting navigators in obtaining SA Level 3.

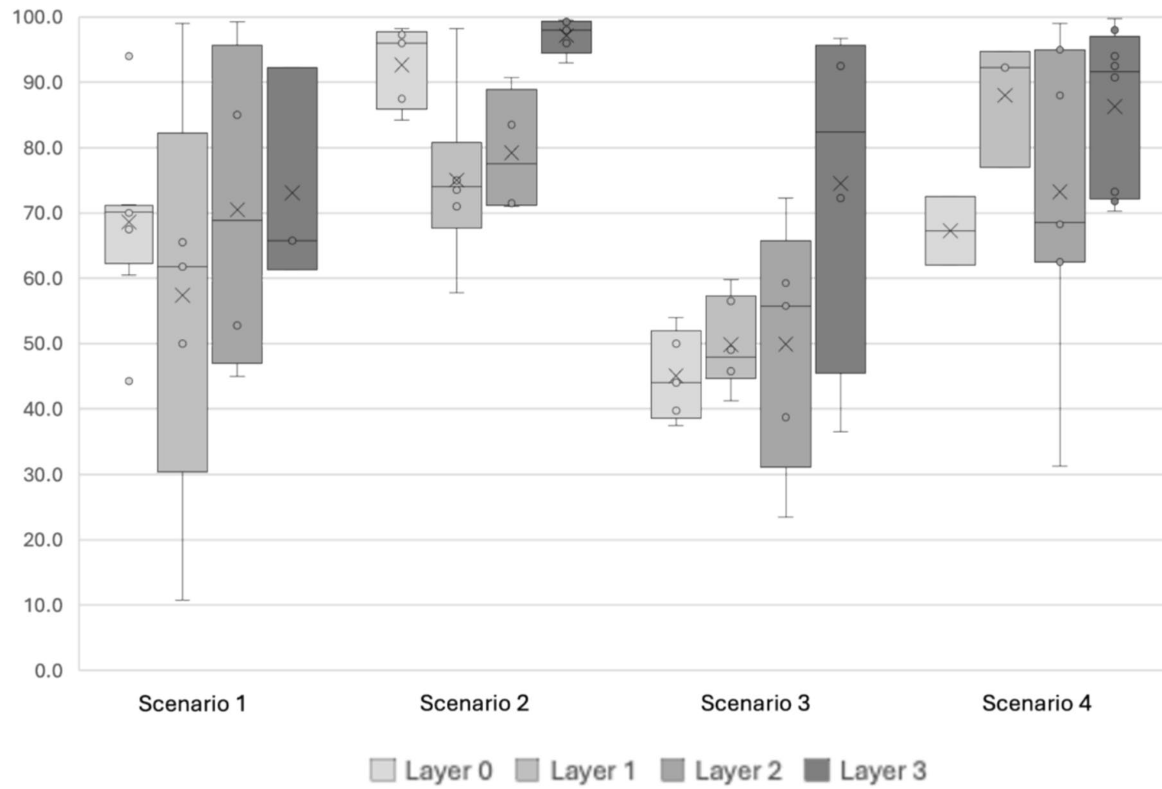


Fig. 11 Box plot for SA Level 3 scores for participants with each layer within each scenario. Error bars represent the 95% confidence interval

4.2 RQ2: Which layer is the most effective?

To evaluate the effectiveness of the transparency layers and thereby the performance of the participants, we measured correct (true) or incorrect (false) decisions, which were then categorized based on the nature of the action taken—"positive" for active decisions and "negative" for passive decisions, as described in the Methods section.

In each scenario, there was a correct decision, which was either to intervene or not intervene. In Scenarios 1 and 3, the correct decision was to intervene with the autonomous CAS decision, and in Scenarios 2 and 3, the correct decision was to not intervene.

The graphs in Fig. 13 summarize the mean scores for each transparency layer and scenario, where intervening correctly = 1 and incorrectly = 0.

The mean values for correct interventions per scenario and layer were between 0.6 and 0.7. Through all scenarios, the mean score for Layer 3 appears to be the layer that obtained the most correct interventions with the system's decision. As illustrated in Fig. 13, a large variance was found in performance between the scenarios and layers (Fig. 14).

A statistically significant difference was not found between the utilized transparency layers and the correct decision.

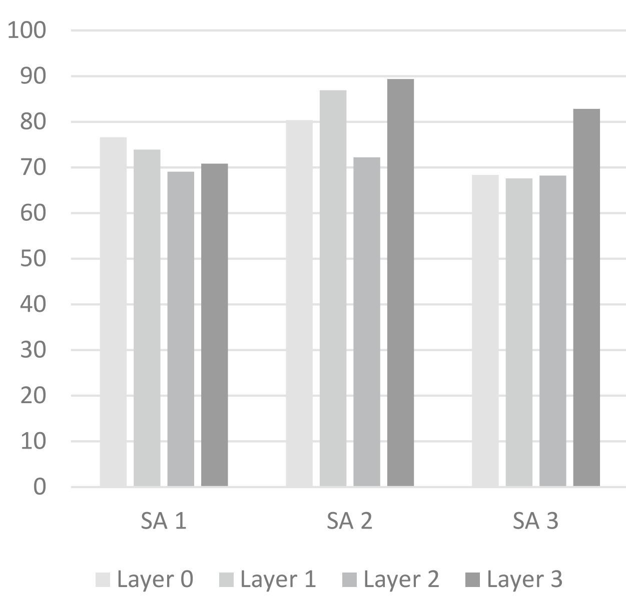


Fig. 12 The mean SA scores through all scenarios per layer

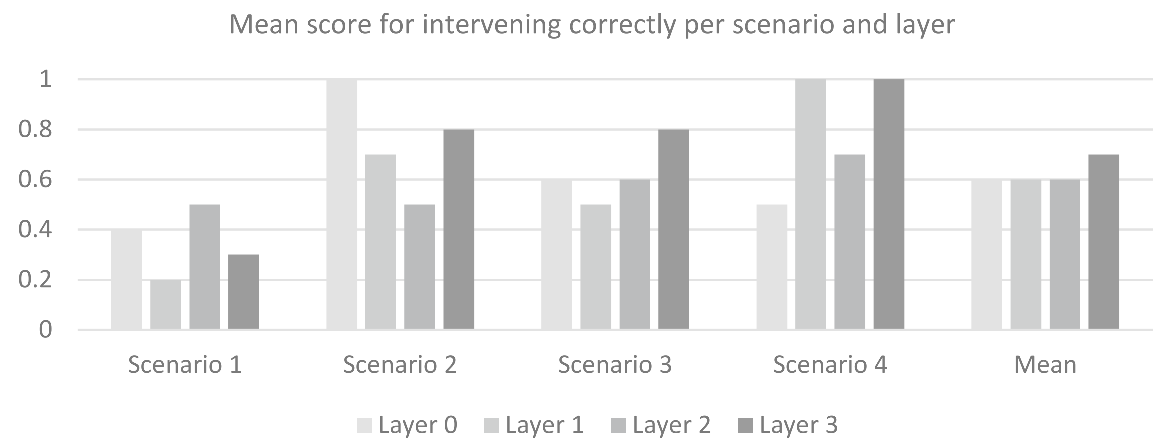


Fig. 13 Mean values for correct interventions for each layer in each scenario. 1 = correct and 0 = incorrect

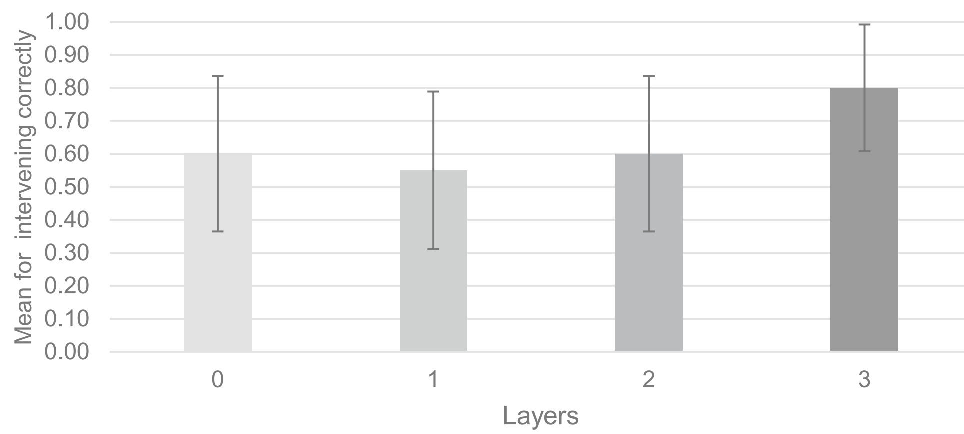


Fig. 14 Histogram with the mean score for intervening correctly by layer. Error bars represent the 95% confidence interval

There were two distinct types of correct behavior in these operations: (1) intervening when needed and (2) not intervening when not needed, referred to as true positives and true negatives, respectively. Similarly, there were two types of errors: (1) intervening when not needed and (2) failing to intervene when needed. The two types of errors are referred to as Type 1 and Type 2 errors, or false positives and false negatives, respectively. For humans, there might be a difference in deciding to actively intervene with the AI's decision and in making the decision to accept the AI's decision. To understand this, the data were counted in a confusion matrix (see Table 4).

Table 4 Confusion matrix counting interventions in the first column and no interventions in the second column

		Performance by participants	
		Positive (Active)	Negative (Passive)
Correct performance	Positive (Intervention required)	True positive 19	False negative 21
	Negative (Intervention not required)	False positive 8	True negative 32

As mentioned, in Scenarios 1 and 3, the correct action was to intervene, and any intervention was considered a *true positive* action, while failure to intervene was a *false negative* action. For Scenarios 2 and 4, the correct action was to not intervene, thus a decision not to intervene was a *true negative* action, while a decision to intervene was a *false positive* action.

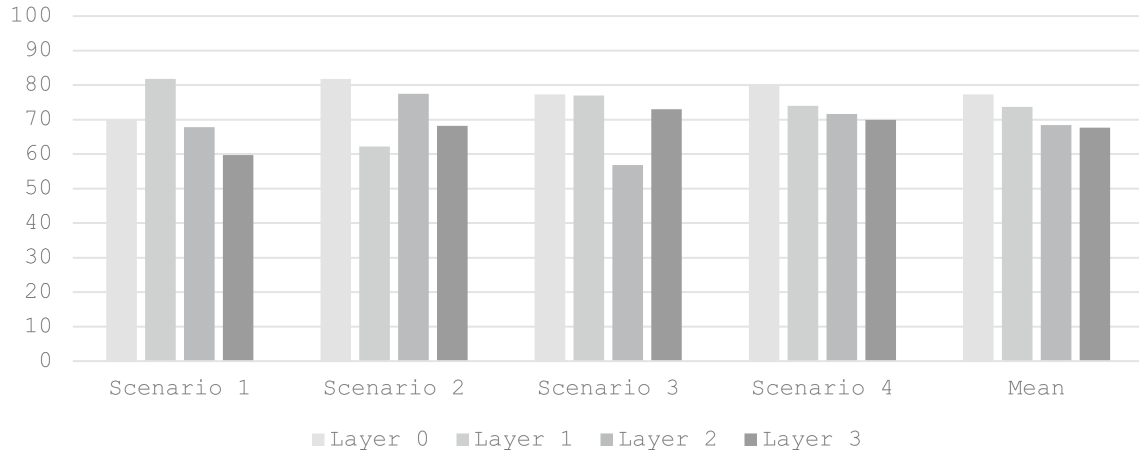


Fig. 15 The mean value of correct interventions and mean value of certitude per scenario and transparency layer

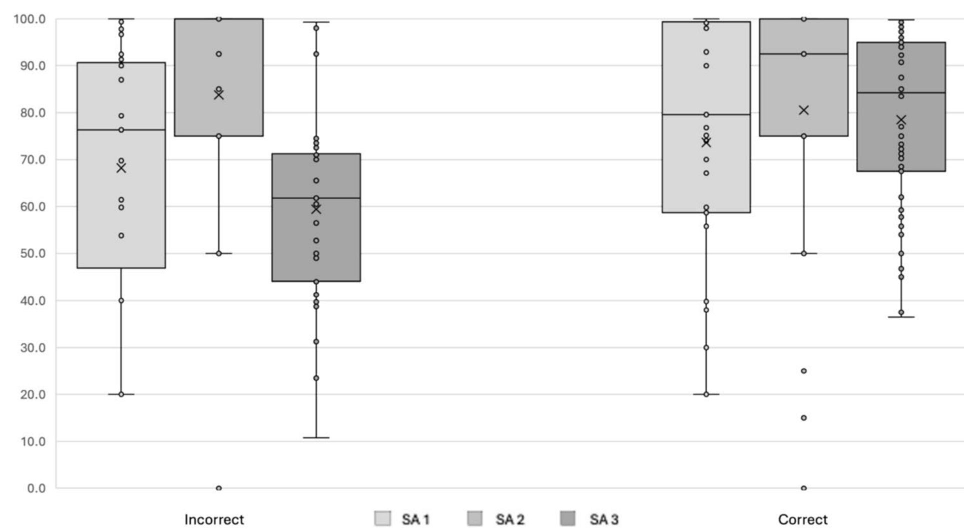


Fig. 16 Incorrect ($n=29$) and correct interventions ($n=51$) compared with SA levels

From Table 4, we can see a higher count of true negative actions ($n=32$) than true positive actions ($n=19$), resulting in a higher count of false negatives ($n=21$) and a lower count of false positives ($n=8$). This means that participants often failed to intervene with the system when they should have. All participants performing positive decisions were asked to justify why they intervened with the system. For those who performed a false positive decision, one did so due to uncertainty or a lack of trust in another vessel's maneuver, two wanted to reduce speed instead of change course, one intervened to start the maneuver sooner, and four misunderstood the situation. Of the participants who misunderstood the situation, three of the misunderstandings would probably have been solved with access to transparency layer 3. The fourth had access to this layer but still wanted to maintain course and speed in Scenario 2. The maneuver would probably not lead to a collision, but the ownship

Table 5 The participants preference of layers from the most preferred to the least preferred

Layer	1. Preference	2. Preference	3. Preference	4. Preference
0	1	1	1	17
1	0	4	13	3
2	1	15	5	0
3	18	0	1	0

would pass very close to the crossing ferry's stern and not be in line with good seamanship.

Furthermore, the participants were asked to rate how certain they were of their understanding of the situation on a linear scale of variable *certitude*, with 0 = uncertain and 100 = certain (Fig. 15).

The results indicate that certitude among the participants decreased when they had a higher transparency layer. In all scenarios, the certitude was lower when participants had access to Transparency Layer 3 than when they had access to Layer 0. No correlation was found between performance (correct actions) and certitude. However, the data strongly indicates that certitude was correlated with how tall the participants were.

SA serves as a foundation for decision-making; therefore, the decision to intervene was compared to the SA score (Fig. 16).

In situations in which participants intervened with the system's decisions, a higher SA Level 3 score was seen.

4.3 RQ3: Which layer provides the best user satisfaction? (rate from 1 to 4)

To answer RQ3, we utilized a card-ranking technique in which an illustration of each transparency layer was displayed on different cards. The participants were then asked to put the cards in order from the one they preferred the most to the one they preferred the least.

Table 5 shows the participants preferred layers. The most preferred transparency layer was Layer 3, and the results steadily decreased to Layer 0, which was the least preferred. This indicates that the participants preferred more information and found this useful. By following this logic, we see that some participants (n = 4) preferred Layer 1 (bounding box) over Layer 2 (give-way situation).

5 Discussion

This section discusses the results for each RQ through a combined discussion of the results.

5.1 Situation awareness (SA)

The access to transparency layers did not result in improved Level 1 SA for the navigators, and the navigators' Level 1 SA seemed to decrease when transparency layers were added. This might be reasonable, as the extra information associated with the transparency layers directs the navigator's attention toward other aspects of SA than merely Level 1 SA.

When examining the results for SA Level 2, Transparency Layer 2, which represented the autonomous CAS's Level 2 SA, provided the overall lowest score for the navigator's Level 2 SA. Layer 2 did not seem to provide satisfactory Level 2 SA (only 72.2), while Layer 3 provided the overall highest rating. The reason for the low effect of Transparency Layer 2 might be that it is always correct in the tested scenarios (yellow sign displaying: give-way situation) and the solution may appear to be given, thus drawing the participant's attention away from other aspects. If there had been a discrepancy between the CAS's decision and SA, the effect on the navigators might be different than that observed here. However, since the construct of human SA is complex, the process of gaining SA for the human may not necessarily be as linear as for the CAS, and access to one of the system's SA levels cannot be expected to align with improved human SA at the same level.

For Level 3 SA, access to transparency layer 3 provided a relatively high score for the navigators across all scenarios. This transparency layer is much better than the others at supporting navigators in obtaining Level 3 SA. The opportunity for the participants to fast-forward both the projection of the autonomous CAS decision and the projections of how the system believes the near future will unfold if maintaining course and speed appears to strengthen the navigator's Level 3 SA. Acquiring SA Level 3 requires an overall understanding that favors a complete and systematic presentation of information.

The results suggest varying effects across different transparency layers and SA levels. Transparency layers may influence attention allocation and decision-making processes, potentially affecting SA levels differently. Conversely, in the SA Level 3 SA projection, transparency layer 3 consistently yielded relatively high scores across all scenarios, suggesting that the layer enhances SA Level 3. In the data, it is difficult to determine whether this is due to the fast-forward function alone or is a product of all transparency layers. However, when comparing the mean SA score across all SA levels for those utilizing transparency layer 3 with the mean SA score for those that did not, the results show that access to transparency layer 3 provided significantly higher scores for SA Levels 2 and 3. For SA Level 1, access to transparency layer 3 did not improve the navigator's score; rather, it was slightly worse. Perhaps Transparency Layer 3 for Level 1 SA complicated something simple. However, transparency layer 3 was found to be better at SA Level 2 than the other layers. This may be because this layer helps navigators gather information (in a somewhat simple process). Overall, the navigators had better scores for SA Level 2 than for Level 3, which was not unexpected. It is reasonable that the answers to the queries were worse for SA Level 3 than SA Level 2, which is a simpler process for the navigator. For something as difficult as projecting how the situation will evolve in the immediate future, a good understanding is needed.

The transparency layers provided foundational support for situational awareness but may fall short when dealing with the inherent complexity and opacity of some AI-driven decisions. In such cases explainability—where users understand the 'why' behind a CAS decision—could be more beneficial than transparency alone.

5.2 Effectiveness

Based on an evaluation of the effectiveness of transparency layers in decision-making across scenarios, Transparency Layer 3 led to the highest number of correct interventions with the system's decision, but without statistical significance. The results also show that those who intervened correctly generally had higher Level 3 SA than those who did not. While potential skewness in scenario complexity or varying access to transparency layers might influence these results, they suggest a connection between robust SA and correct decisions. With the link between SA and effectiveness, and with the result finding that transparency Layer 3 significantly improved the navigators' SA Level 3, it is reasonable to focus on harnessing and refining this layer to maximize its efficiency, making the system more compatible with the navigators' needs. However, a large spread in both performance and SA was observed across all scenarios in all layers. Based on the data, it is difficult to determine why this was the case. One hypothesis is that this was caused by individual differences in skills, as some navigators are simply better at the task than others, independent from the transparency layers. Another hypothesis is that the familiarization time was short. If the participants were experienced with the CAS, they would perhaps understand the system and the transparency layers better, thereby being more capable of intervening when needed.

Furthermore, a closer examination revealed an interesting aspect regarding the navigators' decision-making tendencies: the distinction between actively (positive) intervening in the AI's decision and passively (negative) accepting it. The confusion matrix highlights a higher frequency of true negative actions, indicating a tendency to refrain from intervention when necessary, which suggests the challenge of actively intervening with the system. The data indicate that it is much easier to decide when the correct action is "do nothing" (true negative) than to decide to actively intervene with the system (true positive). Only eight times did the participants have false positive actions. Naturally, we argue that, although unnecessary interventions could potentially lead to lower efficiency and indicate low trust in the capability of the system, the main goal should be to reduce false negatives, which is a greater danger to safety. The autonomous CAS did not display traditional collision avoidance attributes, such as the closest point for approach (CPA), time to CPA, or bow crossing range, which meant that the navigators had to seek this information in the radar if needed. Including this information in Transparency Layer 3 might have reduced the number of false negatives, as actual proximity might be difficult to determine in an electronic chart due to differences in scale.

The results also show that the participants' confidence in their understanding of the situation decreased with higher transparency layers (Fig. 15). This might be a result of information overload, as higher transparency layers introduce more information and may overwhelm users, reducing their confidence in their understanding. However, this was not reported by the users, and the higher transparency layers were found to have the best user satisfaction. Another possible explanation may be that the transparency layers represent increased cognitive load since the navigators are not familiar enough with the system. The explanation for this phenomenon is not clear; it might even be a result of an effect in which access to more information makes the navigators reflect on their initial understanding of the situation, thereby reporting less certitude.

The results also show a significant link between the participants' reported certitude and physical height, with taller participants reporting generally higher certitude. This is reasonable, as height is a known proxy for self-confidence [35]. This result is interesting for those who are both developing and approving systems for collision avoidance, as it is a reminder that simply asking the participants how certain they are of their own understanding is not sufficient when measuring the effect of a particular system.

While the transparency layers in this study provided structured layers of information based on the three-level model of SA, other approaches, such as natural language explanations, may enhance navigators' understanding by providing contextual or scenario-specific guidance. Hodne et al. [36] tested a Conversational User Interface (CUI) for ship to ship communication, but found the CUI to be less trustworthy than communicating with a human ship officer. Still, natural language explanations could support navigators in interpreting the CAS decision in plain terms. Transparency layer 3 is an interactive visualizations which allowed the users to explore 'what-if' scenarios, and is indeed the most effectfull of these layers. Future work may explore alternatives to provide a more holistic understanding of transparency design in autonomous systems.

5.3 Satisfaction

The card-ranking technique revealed that participants preferred transparency layer 3 the most and Layer 0 the least, suggesting a preference for increased information. Some participants, particularly those with extensive maritime experience, favored layer 1 over layer 2, potentially due to perceiving layer 2's information as redundant. This may be an indication that they did not find this information useful since the decision of the AI-based system is to give way (indicated by Layer 0). It may be seen as overabundant information to these participants. Three of these participants were over 50 years old and had more than 10 years of experience at sea. The result for Layer 2 was expected, but in a situation with a discrepancy between the AI-based systems' decision and comprehension, for example, when the system says that the comprehension is to stand on when it should give way or in the case of any other inconsistencies in logic between the system's decision and comprehension, the user might find this transparency layer useful.

Layer 2 might still be useful in scenarios in which discrepancies between the AI-based system's decisions and comprehension occur. In an HCD approach, effective design improves safety, but there are challenges to consider. While users should be involved, they should not be seen as codesigners because they may not fully grasp their own needs and could make erroneous assumptions. Failing to account for this might result in overly complex designs [37]. Although the participants in this study preferred having transparency layers compared to having none, this is not necessarily a confirmation that the content and design of each layer are correct in its current form, but rather an indication that decision transparency has been approached correctly.

6 Conclusions

In this study, we explored how the introduction of transparency layers on an autonomous CAS's decisions influenced a human navigator's individual SA in a full-mission simulator. Our findings across different RQs reveal nuanced insights into the effect between transparency layers, SA levels, and decision-making effectiveness.

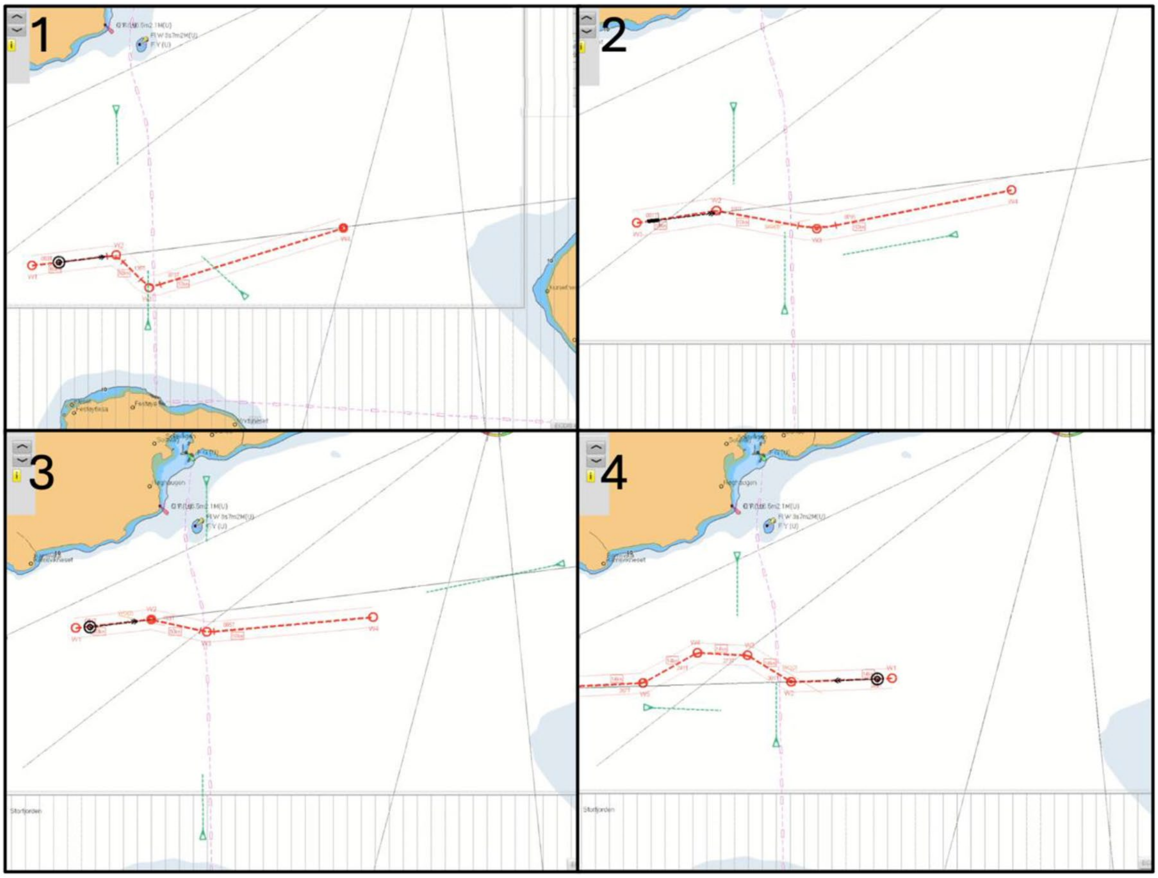
For RQ1, regarding SA, our results indicate that the introduction of transparency layers does not uniformly enhance SA across all levels. Specifically, Level 1 SA appeared to decrease with added transparency layers, suggesting that the additional information might redirect attention away from fundamental SA elements. Level 2 SA was lowest with transparency layer 2, potentially because its consistent correctness would lead participants to overlook other critical aspects of the scenarios. In contrast, transparency layer 3 enhanced Level 3 SA. However, a large variance was observed among the participants that must be considered in future work.

RQ2 focused on the effectiveness of interventions based on the navigators' decisions. Here, transparency layer 3 stood out by leading to the highest number of correct interventions, although without statistical significance. Those making correct interventions generally exhibited higher SA Level 3 scores. The data also highlighted a tendency toward passivity in decision-making, with a higher occurrence of true negative actions, underscoring the challenge of actively engaging with the system.

RQ3 examined user satisfaction, revealing a clear preference for transparency layer 3 among the participants. This preference underscores the importance of providing comprehensive and relevant information to support decision-making. However, the preference for increased information does not necessarily validate the current design of each layer, suggesting room for refinement to better align with user needs and cognitive capabilities.

In conclusion, while transparency layers, particularly Layer 3, show promise in enhancing SA and decision-making effectiveness, the complexity and potential for information overload warrant a careful approach. Designers must balance the desire for transparency with the cognitive limitations of users, ensuring that systems support effective decision-making without overwhelming users. The responses to the different transparency layers highlight the importance of tailoring information presentation to user needs and the specific context of decision-making tasks. These findings underline the need for structured transparency in autonomous systems, suggesting that systems like CAS benefit from transparency layers. By evaluating the effect of detailed transparency on SA and performance, this research informs future AI system designs. Further research should focus on refining these layers to optimize their effectiveness and user satisfaction within HCD frameworks.

Appendix



Scenario 1, The southbound ship should give way for the ownship, and the ownship should give way for the northbound ship. The ship heading west-northwest does not appear to be a risk of collision at this stage. Either way, the ownship should give way to this ship as well since this is a crossing situation. The system's decision appears to consider this and gives way to both ships approaching from the starboard side. But as the scenario evolves, the ship heading west-northwest changes its course towards the west, creating a head-on situation. Thus, the maneuver decision by the system is not sufficient, and the turn towards starboard must start earlier, meaning that the correct action for the navigator would be to intervene.

Scenario 2, The participants are in a crossing situation, where they should give way to the target on the starboard side (northbound target). There is also a target approaching on the port side, which should give way to the ownship. There is no risk of collision with the westbound ship, even though they pass each other starboard to starboard, which can be considered unorthodox. The system's decision complies with the COLREG and passes astern of the northbound target. The correct action is therefore to not intervene.

Scenario 3, The southbound ship should give way to the ownship, and it is not identified as a risk of collision with the northbound ship. The westbound ship is approaching head-on, meaning that both ships should make a maneuver to starboard. In this scenario, the westbound ship is a large cruise vessel. And as the scenario develops, both coastal ferries give way to the cruise ship. This is custom in this area and often agreed upon between the pilot and the respective captains via VHF radio. An indication that this is happening, is that the southbound ferry is on the east side of the rocks, which deviates from its normal sailing route. The result of this is that the system starboard maneuver is not sufficient, it will conflict with the southbound ferry's maneuver, and the cruise ship is restricted from turning starboard since the ferry is giving way. Intervene, start earlier, and turn more to starboard to increase the passing distance with the cruise ship.

Scenario 4, The northbound ferry should give way to the ownship, the ownship should give way to the southbound ferry, and there is no risk of collision with the eastbound ship. The system decision is a starboard maneuver, giving way to the southbound ferry. By examining the ship's vector, we can see that the northbound ferry will pass ahead of the ownship, with a safe distance. The correct action is not to intervene.

Acknowledgements This study is conducted as part of the Center for Research-based Innovation AutoShip and the SAFEMATE project, with financial support from the Research Council of Norway (project number 327903 and 309230 respectively)

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital).

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Aibn, Daibn (2019) Marine 2019/08 part one report on the collision on 8 November 2018 between The Frigate Hnoms Helge Ingstad and the oil Tanker Sola Ts Outside The Sture Terminal In The Hjeltefjord In Hordaland County. https://havarikommisjonen.no/Sjofart/Avgitte-rapporter/2019-08-eng
2. IMO (2021) Outcome of the regulatory scoping exercise for the use of maritime autonomous surface ships (MASS) MSC.1/Circ.1638. https://wwwcdn.imo.org/localresources/en/MediaCentre/PressBriefings/Documents/MSC.1-Circ.1638 - Outcome of the regulatory scoping exercisefor the use of maritime autonomous surface ships... (Secretariat).pdf
3. Brandsæter A, Madsen AN (2024) A simulator-based approach for testing and assessing human supervised autonomous ship navigation. J Marine Sci Technol. https://doi.org/10.1007/s00773-024-00998-1
4. Madsen AN, Brandsæter A, Aarset MV (2023) Decision Transparency for enhanced human-machine collaboration for autonomous ships. Hum Fact Robots Drones UnmannSyst 93:76–84. https://doi.org/10.54941/ahfe1003750
5. Endsley MR (2023) Ironies of artificial intelligence. In Ergonomics (Vol. 66, Issue 11, pp. 1656–1668). https://doi.org/10.1080/00140139.2023.2243404
6. Endsley MR (2017) From here to autonomy: lessons learned from human-automation research. Hum Fact 59(1):5–27. https://doi.org/10.1177/0018720816681350/ASSET/IMAGES/LARGE/10.1177_0018720816681350-FIG4.JPEG
7. van de Merwe K, Mallam S, Nazir S (2022) Agent transparency, situation awareness, mental workload, and operator performance: a systematic literature review. Hum Fact. https://doi.org/10.1177/00187208221077804
8. van de Merwe K, Mallam S, Nazir S, Engelhardtsen Ø (2024) The Influence of Agent Transparency and Complexity on Situation Awareness, Mental Workload, and Task Performance. Journal of Cognitive Engineering and Decision Making. https://doi.org/10.1177/15553434241240553
9. ISO (2019) Part 210: Human-centred design for interactive systems. In ISO 9241–210:2019 Ergonomics of human-system interaction (2nd ed., pp. 1–33). International Standard
10. Norman DA (2013) The design of everyday things. In Interactions (Revised an). Basic Books. https://doi.org/10.1145/1340961.1340979
11. Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Fact 37(1):32–64. https://doi.org/10.1518/001872095779049543
12. Madsen AN (2024) Kim TE (2024) A state-of-the-art review of AI decision transparency for autonomous shipping. J Int Maritime Safety Environ Affairs Shipp. https://doi.org/10.1080/25725084.2024.2336751
13. Madsen AN, Aarset MV, Alsos OA (2022) Safe and efficient maneuvering of a Maritime Autonomous Surface Ship (MASS) during encounters at sea: a novel approach. Maritime Transp Res 3:100077. https://doi.org/10.1016/J.MARTRA.2022.100077
14. Brandsaeter A, Brandsaeter A, Knutsen KE (2018) Towards a framework for assurance of autonomous navigation systems in the maritime industry. https://www.researchgate.net/publication/326020405
15. Brandsaeter A, Smefjell G, Van De Merwe K, Kamsvåg V (2020) Assuring safe implementation of decision support functionality based on data-driven methods for ship navigation. https://doi.org/10.3850/978-981-14-8593-0
16. Asyali E (2012) The Role of ECDIS on Improving Situation Awareness". 13th Annual General Assembly of the IAMU, 15, 17–27. https://www.researchgate.net/publication/320305464
17. Kaasa S (1976) Navigasjonsinstrumenter 2 (13th ed.). Aschehoug
18. Grech MR, Horberry T, Smith A (2002) Human error in maritime operations: analyses of accident reports using the leximancer tool. Proc Hum Fact Ergon Soc Annual Meet 46(19):1718–1721. https://doi.org/10.1177/154193120204601906
19. MAIB (2004) Bridge Watchkeeping Safety Study
20. Maternová A, Materna M, Dávid A, Török A, Švábová L (2023) Human error analysis and fatality prediction in maritime accidents. J Marine Sci Eng 11(12):2287. https://doi.org/10.3390/JMSE11122287
21. Merat N, Seppelt B, Louw T, Engström J, Lee JD, Johansson E, Green CA, Katazaki S, Monk C, Itoh M, McGehee D, Sunda T, Unoura K, Victor T, Schieben A, Keinath A (2019) The "Out-of-the-Loop" concept in automated driving: proposed definition, measures and implications. Cogn Technol Work 21(1):87–98. https://doi.org/10.1007/S10111-018-0525-8/FIGURES/3
22. Endsley MR, Kiris EO (1995) The out-of-the-loop performance problem and level of control in automation. Hum Fact J Hum Fact Ergon Soc 37(2):381–394. https://doi.org/10.1518/001872095779064555
23. Gouraud J, Berberian B, Delorme A (2018) Link between out-of-the-loop performance problem and mind wandering: how to keep the operator in the loop. In: Ayaz H, Dehais F (Eds.), Neuroergonomics The Brain at Work and in Everyday Life (Issue

January 2018, p. 239). Academic Press. https://doi.org/10.1016/b978-0-12-811926-6.00054-3
24. Endsley MR, Jones DG (2012) Designing for situation awareness - an approach to user-centered design (2nd ed.). Taylor & Francis Group, LLC
25. Wickens CD (2015) Situation awareness: its applications value and its fuzzy dichotomies. J Cogn Eng Dec Mak 9(1):90–94. https://doi.org/10.1177/1555343414564571
26. Endsley MR, Jones WM (2001) A model of inter and intra-team situation awareness: implications for design, training and measurement. In: McNeese M, Salsa E, Endsley M (Eds.), New trends in cooperative activities: understanding system dynamics in complex environments (pp. 1–24). Human Factors and Ergonomics Society
27. Stanton NA, Stewart R, Harris D, Houghton RJ, Baber C, McMaster R, Houghtonx RJ, Baberx C, McMasterx R, Salmon P, Hoyle G, Walker G, Young MS, Linsell M, Dymott R, Green D (2006) Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. Ergonomics 49:1288–1311. https://doi.org/10.1080/00140130600612762
28. Barredo Arrieta A, Díaz-Rodríguez N, Ser JD, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrera F (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inform Fus 58:82–115. https://doi.org/10.1016/j.inffus.2019.12.012
29. Druce J, Niehaus J, Moody V, Jensen D, Littman ML (2021) Brittle AI, Causal confusion, and bad mental models: challenges and successes in the XAI Program. https://arxiv.org/abs/2106.05506v1
30. Rutledal D, Relling T, Resnes T (2020) It's not all about the COLREGs: A case-based risk study for autonomous coastal ferries. IOP Confer Series Mater Sci Eng. https://doi.org/10.1088/1757-899X/929/1/012016
31. Endsley MR (1988) Situation awareness global assessment technique (SAGAT). Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, 789–795. https://doi.org/10.1109/NAECON.1988.195097
32. Endsley MR (2000) Direct measurement of situation awareness: Validity and use of SAGAT. In: Endsley MR, Garland DJ (Eds.), Situation awareness analysis and measurement (pp. 147–174). Lawrence Erlbaum Associates. https://doi.org/10.4324/9781315087924-9
33. Ting KM (2010) Confusion Matrix. In: Sammut C, Webb GI (Eds.), Encyclopedia of Machine Learning (p. 209). Springer US. https://doi.org/10.1007/978-0-387-30164-8_157
34. Fincher S, Tenenberg J (2005) Making sense of card sorting data. Expert Syst 22:89–93. https://doi.org/10.1111/j.1468-0394.2005.00299.x
35. Judge TA, Cable DM (2004) The effect of physical height on workplace success and income: preliminary test of a theoretical model. J Appl Psychol 89(3):428–441. https://doi.org/10.1037/0021-9010.89.3.428
36. Hodne P, Skåden OK, Alsos OA, Madsen A, Porathe T (2024) Conversational user interfaces for maritime autonomous surface ships. Ocean Eng 310:118641. https://doi.org/10.1016/J.OCEANENG.2024.118641
37. Lützhöft M, Vu VD (2018) Design for safety. In: Oltedal H, Lützhöft M (Eds.), Managing maritime safety (1st ed., pp. 106–140). Routledge. https://doi.org/10.4324/9780203712979-8